Amendments to the Claims:

This listing of claims will replace all prior versions, and listing, of claims in the application:

Listing of Claims:

Claim 1 (currently amended): A computer-implemented method for predicting one or more locations of single nucleotide polymorphisms in a nucleic acid sequence, comprising the steps of:

calculating a variation frequency from a first base to a second base within a group of bases in a dataset of two or more genes;

generating a variation predictiveness matrix from the calculated variation frequency for each first base to each second base;

comparing the nucleic acid sequence which is not obtained from the dataset of two or more genes, one or more bases in the nucleic acid sequence at a time, with the variation predictiveness matrix to assign a variation value to the one or more bases in the nucleic acid sequence;

identifying the locations of the one or more bases in the nucleic acid sequence where single nucleotide polymorphisms will likely occur based on the assigned variation value; and

outputting the identified locations of the single nucleotide polymorphisms in the nucleic acid sequence where single nucleotide polymorphisms will likely occur to a user via a computer display, an electronic file or a printer.

Claim 2 (previously presented): The method of claim 1, wherein the nucleic acid sequence further comprises one or more chemical modifications.

Claim 3 (previously presented): The method of claim 2, wherein the chemical modifications include methylation or other chemical groups that incorporate additional

charge, polarizability, hydrogen bonding, electrostatic interaction, and fluxionality to one or more bases in the nucleic acid sequence or to the nucleic acid sequence as a whole.

Claim 4 (canceled)

Claim 5 (previously presented): The method of claim 1, wherein a variation from the first base to the second base is nonsynonymous.

Claim 6 (previously presented): The method of claim 1, wherein a variation from the first base to the second base is synonymous.

Claim 7 (original): The method of claim 1, further comprising the step of generating a dataset of single nucleotide polymorphisms for one or more nucleic acid sequences.

Claim 8 (canceled)

Claim 9 (previously presented): The method of claim 1, wherein the dataset comprises genes with nucleic acid chemical modifications.

Claim 10 (previously presented): The method of claim 9, wherein the chemical modifications include methylation or other chemical groups that incorporate additional charge, polarizability, hydrogen bonding, electrostatic interaction, and fluxionality to one or more bases in a nucleic acid sequence of the genes or to the nucleic acid sequence of the genes as a whole.

Claim 11 (previously presented): The method of claim 1, wherein the dataset of two or more genes comprises a known mutation dataset.

Claim 12 (previously presented): The method of claim 1, wherein the dataset of two or more genes comprises a dataset of known diseases.

Claim 13 (previously presented): The method of claim 1, wherein the dataset of two or more genes comprises a dbSNP database.

Claim 14 (previously presented): The method of claim 1, wherein the dataset of two or more genes comprises a non-human mutation database.

Claim 15 (previously presented): The method of claim 1, wherein the dataset of two or more genes comprises a disease-specific database.

Claim16 (previously presented): The method of claim 1, wherein the dataset of two or more genes comprises a non-human disease database.

Claim 17 (previously presented): The method of claim 1, wherein the dataset of two or more genes comprises a HGMD database.

Claim 18 (currently amended): The method of claim 1, wherein the dataset of two or more genes comprises a linkage database.

Claim 19 (currently amended): The method of claim 1, wherein the dataset of two or more genes comprises a splice variant database.

Claim 20 (previously presented): The method of claim 1, wherein the dataset of two or more genes comprises a translocation database.

Claim 21 (previously presented): The method of claim 1, wherein the dataset of two or more genes comprises a database of known mutations.

Claim 22 (previously presented): The method of claim 1, further comprising the step of adjusting the calculated variation frequency for wild type genes.

Claim 23 (previously presented): The method of claim 1, further comprising the step of adjusting the calculated variation frequency for engineered or non-naturally occurring genes.

Claim 24 (previously presented): The method of claim 1, further comprising the step of adjusting the calculated variation frequency for conservative polymorphisms.

Claim 25 (previously presented): The method of claim 1, further comprising the step of adjusting the calculated variation frequency for non-conservative polymorphisms.

Claim 26 (previously presented): The method of claim 1, further comprising the step of adjusting the calculated variation frequency for cDNA stability.

Claim 27 (previously presented): The method of claim 1, further comprising the step of adjusting the calculated variation frequency for predicted DNA structure.

Claim 28 (previously presented): The method of claim 1, further comprising the step of adjusting the calculated variation frequency for predicted RNA structure.

Claim 29 (previously presented): The method of claim 1, further comprising the step of adjusting the calculated variation frequency for predicted protein structure.

Claim 30 (previously presented): The method of claim 1, further comprising the step of adjusting the calculated variation frequency for post-translational modification sequences.

Claim 31 (previously presented): The method of claim 1, further comprising the step of adjusting the calculated variation frequency for protein stability.

Claim 32 (previously presented): The method of claim 1, further comprising the step of adjusting the calculated variation frequency for predicted protein transport.

Claim 33 (previously presented): The method of claim 1, further comprising the step of adjusting the calculated variation frequency for shuffled genes.

Claim 34 (previously presented): The method of claim 1, further comprising the step of adjusting the calculated variation frequency for site-directed mutagenesis genes.

Claim 35 (previously presented): The method of claim 1, further comprising the step of adjusting the calculated variation frequency for methylated sequences.

Claim 36 (previously presented): The method of claim 1, further comprising the step of adjusting the calculated variation frequency for epigenetic variation.

Claim 37 (previously presented): The method of claim 1, wherein the nucleic acid sequence comprises a cDNA sequence.

Claim 38 (previously presented): The method of claim 1, wherein the nucleic acid sequence comprises a genomic sequence.

Claim 39 (previously presented): The method of claim 1, wherein the nucleic acid sequence comprises an intron/exon boundary.

Claim 40 (previously presented): The method of claim 1, wherein the nucleic acid sequence comprises a transcriptional control sequence.

Claim 41 (previously presented): The method of claim 1, wherein the nucleic acid sequence comprises a transport control sequence.

Claim 42 (previously presented): The method of claim 1, wherein the nucleic acid sequence comprises a translational control sequence.

Claim 43 (canceled)

Claim 44 (previously presented): The method of claim 1, wherein the nucleic acid sequence comprises a splicing control sequence.

Claim 45 (previously presented): The method of claim 1, wherein the variation predictiveness matrix correlates the calculated variation frequency to a change of the first base to the second base within the group of bases from one to ten bases at a time.

Claim 46 (previously presented): The method of claim 1, further comprising the step of normalizing the generated variation predictiveness matrix for the codon usage of a target organism.

Claim 47 (previously presented): The method of claim 1, wherein the dataset of two or more genes comprises a mutant gene dataset that comprises all mutant genes in a mutant gene database.

Claim 48 (previously presented): The method of claim 1, wherein the dataset of two or more genes comprises a mutant gene dataset that comprises all mutant genes in a mutant gene database minus the mutant genes of the mutant gene dataset that are known to cause a disease.

Claim 49 (original): The method of claim 1, wherein the nucleic acid sequence comprises an entire genome.

Claim 50 (original): The method of claim 1, wherein the nucleic acid sequence comprises a human genome.

Claim 51 (original): The method of claim 1, wherein the nucleic acid sequence comprises a gene cluster for a target human disease.

Claim 52 (previously presented): The method of claim 1, wherein the dataset of two or more genes comprises a mutant gene dataset that comprises a human mutation database.

Claim 53 (previously presented): The method of claim 1, wherein the steps are effected by a computer program.

Claims 54-55 (canceled)

Claim 56 (previously presented): The method of claim 1, wherein the step of generating the variation predictiveness matrix is performed in silico and the dataset of two or more genes comprises a human mutant database.

Claim 57 (previously presented): The method of claim 1, wherein the step of comparing the nucleic acid sequence, one or more bases in the nucleic acid sequence at a time, with

the variation predictiveness matrix to assign a variation value to the one or more bases in the nucleic acid sequence is performed in silico.

Claims 58-202 (canceled)

Claim 203 (currently amended): A computer readable medium encoded with a computer program executable by a processor for predicting one or more locations of variations in a wild-type gene sequence, comprising:

a code segment for calculating a variation frequency from a first base to a second base within a group of bases in a nucleic acid dataset;

a code segment for generating a variation predictiveness matrix from the calculated variation frequency for each first base to each second base;

a code segment for comparing the wild-type gene sequence which is not obtained from the dataset of two or more genes, one or more bases in the wild-type gene sequence at a time, with the variation predictiveness matrix to assign a variation value to the one or more bases in the wild-type gene sequence; and

a code segment for identifying one or more locations where a variation is likely to occur in one or more bases of the wild-type gene sequence based on the assigned variation value.

Claim 204 (currently amended): A computer readable medium encoded with a computer program executable by a processor for predicting one or more locations where a variation is likely to occur in one or more codons in a wild-type gene sequence, comprising:

a code segment for calculating a variation frequency from a first codon to a second codon in a mutant gene dataset;

a code segment for generating a codon mutation predictiveness matrix from the calculated variation frequency for each first codon to each second codon;

a code segment for comparing the wild-type gene sequence which is not obtained from the dataset of two or more genes, one or more codons in the wild-type gene

sequence at a time, with the codon mutation predictiveness matrix to assign a variation value to the one or more codons in the wild-type gene sequence; and

a code segment for identifying the one or more locations where the variation is likely to occur in the one or more codons in the wild-type gene sequence based on the assigned variation value.

Claims 205-213 (canceled)